



# Trabajo Fin de Grado

Test clínicos para la detección de fracturas.  
Revisión sistemática  
Clinical tests for fractures detection.  
A systematic review

Autor/es

Aitor Llamazares Alonso

Director/es

Miguel Malo Urriés

Facultad Ciencias de la Salud

2017/2018

## ÍNDICE

<b>RESUMEN</b> .....	3
<b>INTRODUCCIÓN</b> .....	4
<i>Tabla 1</i> .....	7
<b>OBJETIVOS</b> .....	8
<b>METODOLOGÍA</b> .....	9
<i>Diagrama 1</i> .....	12
<i>Figura 1</i> .....	13
<i>Tabla 2</i> .....	14
<b>RESULTADOS</b> .....	16
<i>Tabla 3: Escafoides.</i> .....	17
<i>Tabla 4: Mandíbula.</i> .....	18
<i>Tabla 5: Columna.</i> .....	19
<i>Tabla 6: Pelvis.</i> .....	19
<i>Tabla 7: Codo.</i> .....	20
<i>Tabla 8: Diapasón o percusión con auscultación.</i> .....	21
<b>DISCUSIÓN</b> .....	22
Limitaciones del estudio .....	36
Sugerencias para futuros estudios .....	36
Implicaciones clínicas .....	36
<b>CONCLUSIONES</b> .....	39
<b>PALABRAS CLAVE</b> .....	41
<b>DICCIONARIO</b> .....	411
<b>BIBLIOGRAFÍA</b> .....	411

## **RESUMEN**

### **OBJETIVO:**

El objetivo de la presente revisión sistemática es resumir y actualizar el conocimiento acerca de los test clínicos para detectar fracturas, y determinar que test clínico posee fiabilidad y validez diagnóstica para poder ser utilizado en la clínica.

### **MÉTODOS:**

Una búsqueda bibliográfica fue llevada a cabo en febrero del 2018 en la base de datos Pubmed. La revisión sistemática se realizó siguiendo las normas PRISMA. La calidad metodológica se llevó a cabo con la herramienta QUADAS-2. Un total de 35 artículos cumplieron los criterios de inclusión/exclusión y fueron incluidos en la revisión.

### **RESULTADOS:**

Los resultados fueron analizados en subgrupos, en función de la localización de la fractura o del método utilizado para detectarla. Se forman 6 subgrupos con un número variable de estudios: escafoides, mandíbula, columna, pelvis, codo y diapasón/percusión. Los resultados se plasmaron en tablas con los parámetros: sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo.

### **CONCLUSIONES:**

Se encontraron varios test con validez diagnóstica estadísticamente significativa como para aplicarlos en la clínica. Se trata de los test para diagnosticar las fracturas mandibulares, de codo, cervicales, de escafoides y el test del diapasón. Pese a que los dos últimos subrayan que no pueden ni deben sustituir al diagnóstico por imagen y deben usarse como complemento diagnóstico. No hay evidencia ni validez suficiente para defender el uso de test clínicos en el diagnóstico de las fracturas de pelvis o toraco-lumbares y por lo tanto futuras investigaciones deben centrarse en encontrar una regla de decisión diagnóstica para este tipo de fracturas.

## **INTRODUCCIÓN**

La fractura puede definirse como la interrupción de la continuidad ósea o cartilaginosa.

Hay varias clasificaciones a la hora de hablar de las fracturas. En relación a su etiología, hay varias circunstancias que pueden dar lugar a una fractura. Las fracturas más habituales son las desencadenadas por un traumatismo, su gravedad y pronóstico son directamente proporcionales a la violencia del traumatismo causal. En las fracturas patológicas el factor fundamental es la debilidad ósea. Pueden deberse a procesos generales que cursen con osteopenia u osteosclerosis bien sean enfermedades óseas ó metabólicas. Por último tenemos las fracturas por fatiga o estrés, las cuales se producen por sollicitaciones mecánicas repetidas.

Según el mecanismo de producción tenemos las fracturas por mecanismo directo que se producen en el lugar de impacto de la fuerza responsable y por mecanismo indirecto que se producen a distancia del traumatismo. Estas últimas tienen una clasificación propia:

- Fracturas por compresión: la fuerza actúa en el eje del hueso produciendo un aplastamiento.
- Fracturas por flexión: la fuerza actúa en dirección perpendicular al eje mayor del hueso, produciendo fuerzas de compresión en la concavidad el mismo y fuerzas de distracción en la convexidad.
- Fractura por cizallamiento: El hueso es sometido a una fuerza de dirección paralela y de sentido opuesto, originándose una fractura de trazo horizontal.
- Fracturas por torsión: son el resultado de una fuerza que imprime un movimiento de rotación sobre su eje, estando un extremo fijo o por una rotación en sentido inverso.
- Fracturas por tracción: es el resultado de la acción de dos fuerzas en la misma dirección y sentido opuesto. Los arrancamientos y avulsiones son ejemplo de este tipo de fracturas.

En función de la continuidad ósea las fracturas se pueden dividir en fracturas completas e incompletas. En las fracturas completas existe una solución de continuidad y afecta a todo el espesor del hueso y al periostio. Estas pueden subdividirse en simples, conminutas o con desplazamiento. En las fracturas incompletas; sin embargo, la línea de fractura no abarca todo el espesor del hueso. Aquí podemos encontrar las fisuras, fracturas en callo verde y en caña de bambú.

Por último tenemos las fracturas clasificadas en función de su estabilidad. Las fracturas estables son las que no tienen tendencia a desplazarse tras conseguir la reducción, son fracturas de trazo transversal u oblicuo, menor de 45°. Las fracturas inestables en cambio tienden a desplazarse tras la reducción. (1,2)

Hoy en día las pruebas diagnósticas para detectar fracturas por antonomasia son las pruebas de imagen. Entre ellas tenemos la radiografía, la más común de todas, la resonancia magnética o tomografía computarizada y por ende serán la prueba objetiva o "gold-standard" con la que comparar los test clínicos de los estudios incluidos en esta revisión.

La resonancia magnética (RM), por su alta disponibilidad en atención primaria, utilidad y fácil interpretación debe ser la primera técnica de imagen que se utilice en la patología osteoarticular. La resonancia magnética es una prueba más sensible que la radiología simple para evidenciar múltiples alteraciones intraarticulares de partes blandas. Entre sus inconvenientes se encuentra poseer una baja especificidad en el estudio de patología ósea y el coste elevado de la técnica, que obliga a una indicación individualizada.

La radiología simple; por su parte, no suele ser en sí misma diagnóstica al haber discordancia entre los hallazgos radiológicos y la clínica del paciente.

La tomografía axial computarizada (TAC) es más sensible que la radiología para visualizar calcificaciones, aire y grasa, con el inconveniente conocido de que utiliza radiaciones ionizantes y de forma considerable. (3,4)

Otra consideración que debemos tener en cuenta es la cantidad de radiación ionizada a la que se ven expuestos los pacientes a causa de las imágenes

radiológicas que se les realiza, tanto los rayos-X como los escáneres de tomografía computarizada. (5)

El aumento del coste que supone realizar pruebas por imagen ha dejado en evidencia la necesidad de encontrar una herramienta o un método reproducible que sea efectivo y menos costoso, que ayude al personal sanitario a descartar fracturas cuando existe sospecha de ella en el examen físico inicial. Por lo tanto es interesante este estudio no solo por el ahorro económico que puede suponer a la larga la normalización del uso de los test clínicos validos, sino también por la posibilidad de utilizarlos en medios austeros o con menos recursos como herramienta diagnóstica. (3)

Sin embargo, en la consulta de fisioterapia no se disponen de estos recursos y en ocasiones, tras un traumatismo o dolencia y previo a acudir al hospital los pacientes acuden a los servicios de fisioterapia. Por ello, ante el riesgo de encontrarnos con un paciente que ha sufrido una fractura y que todavía no ha sido diagnosticada, es interesante aplicar test clínicos o establecer exámenes físicos específicos en la exploración física inicial para tratar de detectar una fractura y poder derivar al paciente.

Estos test clínicos deben ser validos, fiables y fácilmente reproducibles para que los pueda aplicar el profesional sanitario sin dificultad. Por ello nuestro estudio, una revisión sistemática de estudios de precisión diagnóstica, se centra en exponer la validez de uno o varios test clínicos, comparándolos con una prueba objetiva o "gold-standard". Los términos de sensibilidad, especificidad y valores predictivos son los más utilizados para referirse al análisis de las propiedades de un test diagnóstico.

- **Sensibilidad:** la sensibilidad caracteriza la capacidad del test clínico para detectar fracturas en sujetos que tienen una fractura.
- **Especificidad:** la especificidad caracteriza la capacidad del test clínico para detectar la ausencia de fractura en sujetos sanos.
- **Valor predictivo positivo (VPP):** probabilidad de tener la fractura si el resultado del test clínico es positivo.
- **Valor predictivo negativo (VPN):** probabilidad de no tener la fractura si el resultado del test clínico es negativo.

Mientras mayor sea la prevalencia de la enfermedad, mayor es el valor predictivo positivo (VPP) y menor es el negativo (VPN) del *test* diagnóstico. Por ello, con el fin de dimensionar el beneficio clínico de un *test* diagnóstico en forma independiente de la prevalencia resulta de mayor utilidad la aplicación de los *likelihood ratio* (LR).

Conocido también como razón de verosimilitud, el *likelihood ratio* (LR) se define como la razón entre la posibilidad de observar un resultado en los pacientes con la enfermedad en cuestión *versus* la posibilidad de ese resultado en pacientes sin la patología. (6,7)

**Tabla 1.** Rangos de valores de *likelihood ratio* y su impacto en utilidad clínica

LR positivo	LR negativo	Utilidad
10	<0,1	Altamente relevante
5-10	0,1-0,2	Buena
2-5	0,5-0,2	Regular
<2	>0,5	Mala

## **OBJETIVOS**

El objetivo principal de la presente revisión sistemática es realizar un análisis de la literatura acerca de la validez diagnóstica de los diferentes test clínicos para detectar fracturas independientemente de donde se hayan producido. Este objetivo consiste en valorar la fiabilidad y la validez diagnóstica de los test clínicos analizados en el estudio. Con estas valoraciones queremos llegar a la conclusión de que test clínicos tienen una buena utilidad clínica como para que los profesionales sanitarios puedan realizarlo y cuáles no; y por lo tanto se deba seguir investigando en ellos o dejar de utilizarlos.

Debido a la variedad de test clínicos en función de la parte del organismo valorada, se agruparán, resumirán y definirán los test clínicos en función de la estructura ósea valorada o la técnica de valoración utilizada.

El objetivo secundario de la revisión será analizar la validez diagnóstica de cada test clínico en función de la región corporal valorada:

- Pie-tobillo.
- Pierna-rodilla.
- Muslo-pelvis.
- Columna lumbo-torácica.
- Columna cervical.
- Región cráneo-facial.
- Escápula-brazo.
- Codo-antebrazo.
- Muñeca-mano.



## **METODOLOGÍA**

### **Bases de datos y estrategia de búsqueda.**

En este estudio se ha seguido la metodología para realizar revisiones sistemáticas propuesta por Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (8). En él se incluye la tabla propuesta por PRISMA para la elección de los artículos incluidos. (Diagrama 1)

Los estudios que se han incluido en esta revisión fueron identificados realizando una búsqueda sistemática a través de la base de datos electrónica Pubmed ya que es la base de datos recoge el mayor número de artículos. Además se consultó con los expertos sobre la materia analizada y se hizo un análisis de la bibliografía de los artículos incluidos en la revisión tras la primera búsqueda. De este modo se ha logrado llegar hasta un gran número de artículos. La búsqueda se realizó obteniendo toda la información relativa al objeto de estudio hasta el 26 de febrero de 2015.

Se utilizó la herramienta thesaurus proporcionada por el propio Pubmed en las siguientes palabras: "Fractures, Bone" (término MeSH), "Mandibular Fractures" (término MeSH), "Sensitivity and Specificity" (término MeSH), "Physical examination" (término MeSH) y "Diagnostic Tests, Routine" (término MeSH). Estos términos MeSH se combinaron con otras palabras genéricas; "fractures", "sensitivity", "specificity", "validity", "reliability", "diagnostic tool", "screening tool", "test", "clinical test", "diagnostic accuracy" y "diagnosis" mediante los marcadores/operadores booleanos "AND" (para la sumación de los términos) y "OR" (si se utiliza más de un término para proponer una idea). De esta manera queda una estrategia de búsqueda planteada en el anexo I.

Tras realizar la búsqueda nos aparecen 822 artículos. No se ponen límites para filtrar la búsqueda ya que queremos que aparezcan el mayor número

de artículos que posteriormente serán filtrados mediante los criterios de inclusión y exclusión.

### **Criterios de elección**

#### **Inclusión:**

- Humanos y predominantemente en poblaciones adultas.
- Se debe delimitar la sensibilidad y especificidad o aportar los datos para extraerlos.
- Deben contener un “gold-standard” o “prueba de referencia” fiable.
- Deben describir el test clínico o examen físico que se lleva a cabo para detectar la fractura.

#### **Exclusión:**

- Estudios con animales.
- Tipos de estudio: revisiones sistemáticas, casos-controles, casos-clínicos y estudios pilotos.
- Estudios en los que el test clínico es un cuestionario.
- Estudios en los que la muestra es exclusivamente niños.
- El test clínico se realiza bajo anestesia.

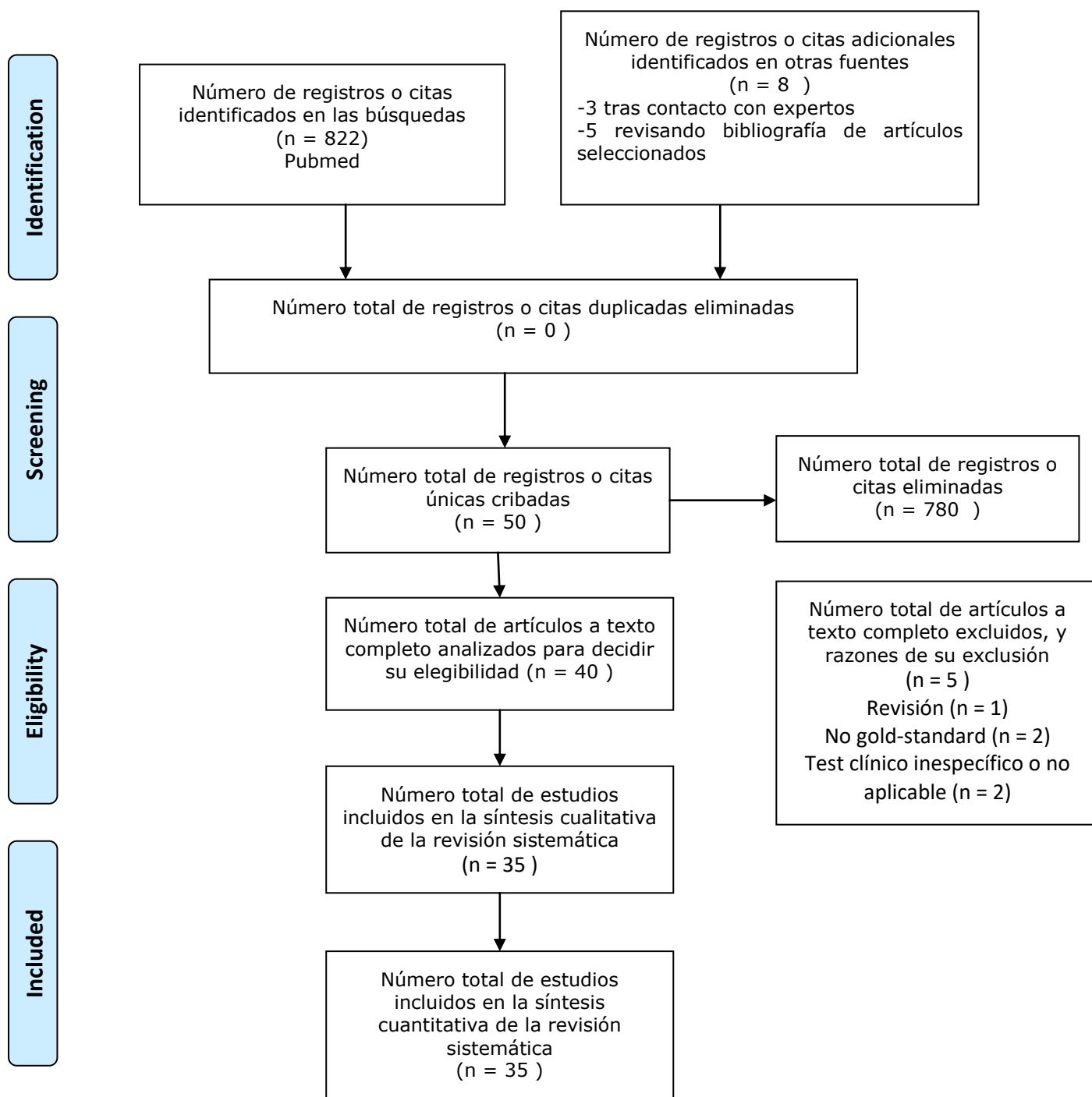
### **Selección del estudio**

En la búsqueda se identificaron 822 artículos potencialmente relevantes para la revisión en cuestión. Tras revisar los títulos y resúmenes, en caso de que el título no fuese suficiente para clarificar si el artículo cumple los criterios de inclusión/exclusión, se descartan 780 artículos.

De los consiguientes 42 artículos se descartan 10 por no poder obtener el texto completo. Nos quedamos con 32 artículos de los cuales descartamos 5 tras la lectura profunda del mismo. En el “anexo II: protocolo de revisión” se muestra el motivo del descarte de estos 5 artículos además de todo el proceso llevado a cabo para completar la selección de los estudios.

Tras analizar los 27 artículos restantes, se revisa la bibliografía de los mismos y se incluyen 5 nuevos artículos. Además tras el contacto con expertos de la materia (Doctor/tutor: Miguel Malo) se incluyen 3 nuevos artículos. El resultado final es de 35 artículos incluidos y analizados en la revisión sistemática.

**Diagrama 1:** selección del estudio



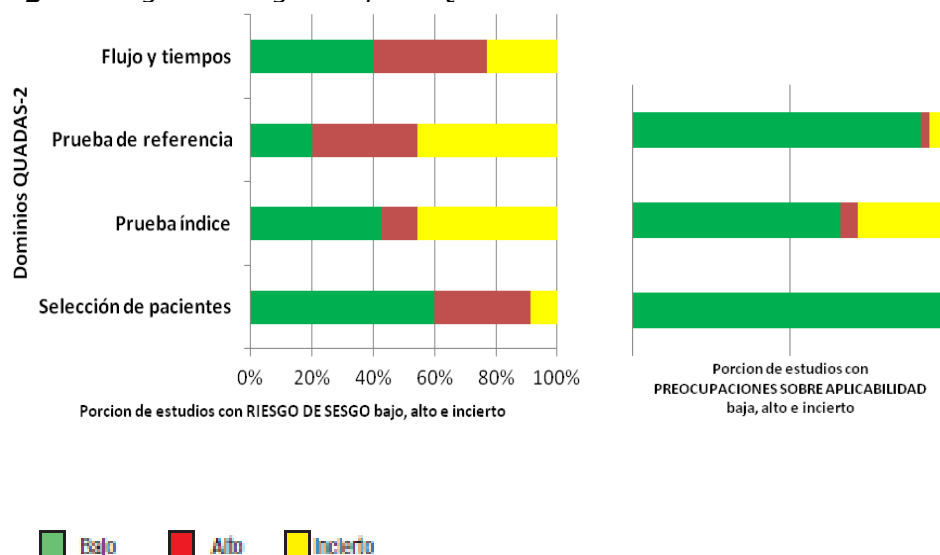
## Evaluación de la calidad metodológica

Las revisiones sistemáticas de estudios de precisión diagnóstica a menudo se caracterizan por resultados marcadamente heterogéneos debido a las diferencias metodológicas de los estudios incluidos, por lo que es fundamental evaluar adecuadamente aspectos que determinan la validez interna y externa de cada uno de ellos, y por ende su calidad.(9)

Para evaluar la calidad metodológica de los estudios incluidos, se utilizó la herramienta QUADAS-2. Esta herramienta consta de cuatro dominios: selección de los pacientes, prueba índice, prueba de referencia, flujo y tiempos. Cada dominio es evaluado en términos de su riesgo de sesgos y los primeros tres dominios también se evalúan por su aplicabilidad. (Figura 1), (Tabla 1).

El riesgo de sesgo se juzga como "bajo", "alto" o "incierto" utilizando las guías elaboradas en la fase 2. Si las respuestas a todas las preguntas orientadoras de un dominio son "sí", entonces la probabilidad de sesgo puede ser juzgada como "baja". Si se contesta "no" a cualquier pregunta orientadora, existe la posibilidad de sesgo y debería calificarse como "alto". La categoría de "incierto" debe utilizarse únicamente cuando los datos son insuficientes para emitir un juicio.(9)

**Figura 1:** gráfica sugerida para QUADAS-2



**Tabla 2:** Presentación tabular sugerida para resultados QUADAS-2.

Estudio	Probabilidad de sesgos				Preocupación sobre la aplicabilidad de los resultados		
	Selección de los individuos	Prueba índice	Prueba de referencia	Flujo y tiempos	Selección de los pacientes	Prueba índice	Prueba de referencia
Caputo et al., 2013	😊	😊	?	😊	😊	😊	😊
Langdon et al., 2010	😊	?	?	😊	😊	😞	😊
Moore, 2009	😊	😊	😞	😊	😊	😞	😊
Tijun et al., 2002	😊	?	😞	?	😊	😊	😊
Rooker et al., 1998	😊	?	😞	😞	😊	😊	😊
Adams y Yarnold, 1997	😞	😊	😊	?	😊	😊	😊
Kazem and Roscoe, 2000	😊	😊	😊	😊	😊	?	?
Alonso LI et al., 1995	😞	😊	😞	😊	😊	😊	😊
Neiner J et al., 2016	😞	😞	?	😞	😊	?	😊
Schwab et al. 1998	😊	?	😞	?	😊	😊	😊
P. T. Grant, 1990	😊	😊	?	😞	😊	?	😊
Siminoski K et al., 2003	?	😊	?	😊	😊	😊	😊
S. J. Ham et al., 1996	😞	?	?	😊	😊	😊	😊
Jawad Z et al., 2012	😞	?	?	😊	😊	😊	😊
Amiri H., et al 2012	😊	😊	😊	?	😊	😊	😊
Appelboom A et al., 2008	😊	😊	😊	😞	😊	😊	😊
Darracq et al., 2008	😞	😊	😊	😊	😊	😊	?
Jie KE et al., 2014	?	?	😞	😊	😊	?	😊
Lamprakis A et al., 2007	?	😊	😊	😊	😊	😊	😊
Walenkamp MM., 2016	😞	😊	?	😞	😊	?	😊
Ghane MR et al., 2016	😞	?	?	😞	😊	😊	😊
Unay K et al., 2009	😞	😞	😞	😞	😊	😊	😊
Rhemrev SJ et al., 2010	😞	😞	😞	😞	😊	?	😊
Hawksworth and Freeland, 1991	😊	?	?	?	😊	😊	😊
Freeland P, 1989	😊	?	😞	😊	😊	😊	😊
Misurva et al., 1987	😞	?	?	?	😊	😊	😊
Hoffman JR et al., 2000	😊	?	?	?	😊	😊	😊
Arundel et al., 2014	😊	😊	?	😞	😊	😊	😊
Fatima ST et al., 2010	😊	?	?	😞	😊	😊	😞
Inaba K et al., 2011	😊	😞	?	😞	😊	?	😊
Inaba K et al., 2016	😊	?	😊	😞	😊	?	😊
Esberger DA, 1994	😊	😊	😞	😞	😊	😊	😊
Waeckerle et al., 1987	😊	?	😞	😊	😊	?	😊
Duane et al., 2007	😊	?	?	?	😊	?	😊
Gonzalez RP et al., 1999	😊	😊	😞	😊	😊	😊	😊

😊 Probabilidad baja. 😞 Probabilidad alta. ? Probabilidad incierta.

Si bien podría considerarse la posibilidad de ser más restrictivo en los criterios de inclusión de la revisión incorporando sólo trabajos de un mínimo de calidad en uno o varios dominios, existe consenso de preferir la inclusión de toda la evidencia disponible, y luego comunicar la calidad de todas las

investigaciones halladas, investigando luego las posibles razones de la heterogeneidad de los hallazgos. (9)

## **Extracción de datos**

La extracción de datos fue realizada utilizando un formulario cedido por el tutor del trabajo. En este formulario se muestran los datos relevantes que se deben sacar de cada estudio para que posteriormente queden plasmados en los resultados. En orden consta de los siguientes apartados:

1. Autor y año.
2. Tipo de estudio.
3. Objetivo del estudio.
4. Muestra.
5. Gold-standard o prueba de referencia.
6. Test clínico aplicado.
7. Resultados. (Se incluyen tablas).
8. Discusión.
  - a. Comentarios/ Discusión del estudio.
  - b. Comparación de resultados con otros (+) y (-).
  - c. Errores propios (Auto-comentados).
  - d. Errores de otros estudios.
  - e. Errores del estudio (comentados en otros estudios).
  - f. Mejoras futuras.
9. Conclusiones.

## **RESULTADOS**

### **Evaluación de la calidad metodológica**

Vistos los resultados, tras la aplicación de la herramienta QUADAS-2 en esta revisión, cabe destacar la cantidad de "incierto" que han aparecido en el riesgo de sesgo, principalmente en el análisis de la "prueba de referencia" (16/35) y en la "prueba índice" (16/35).

El riesgo de sesgo es "alto" en casi un tercio de los artículos en lo que se refiere a "selección de pacientes" (11/35), "prueba de referencia" (12/35) y "flujo y tiempos" (13/35). Esto se puede deber a que una gran cantidad de artículos fueron escritos hace tiempo y no tienen la metodología tan elaborada como en los artículos actuales, por lo tanto muchas preguntas del QUADAS-2 se quedan en el aire sin una respuesta. Además son muchos los artículos en los que se declara que la prueba de referencia no es 100% fiable.

Sin embargo se muestran resultados positivos en cuanto a la preocupación de aplicabilidad debido a que los estudios analizados cumplen con los objetivos de la revisión propuesta. Los resultados son de una preocupación baja en todos los parámetros; "selección de los pacientes" (35/35), "prueba de referencia" (32/35) y "prueba índice" (23/35), pese a que en esta última hay varios "incierto" por falta de información (10/35).

### **Estructuración y organización de los artículos**

Tras realizar el análisis de los 35 artículos incluidos en la revisión, se llegó a la conclusión la necesidad de dividir y organizar los estudios en función del hueso fracturado y testado o en función de la técnica realizada.

Al iniciar la revisión no se quiso filtrar los artículos según la localización de la fractura y centrarnos solo en una ya que hubiese quedado muy pobre en número de artículos. El resultado ha sido de 6 subgrupos dentro de la revisión y cada subgrupo con un número variado de artículos.



Los estudios incluidos en esta revisión son de tipo observacional. Estos estudios, corresponden a diseños de investigación cuyo objetivo es “la observación y registro” de acontecimientos sin intervenir en el curso natural de estos. Las mediciones, se pueden realizar a lo largo del tiempo (estudio longitudinal), ya sea de forma prospectiva o retrospectiva; o de forma única (estudio transversal). En nuestro caso se trata de estudios prospectivos; en los que primero se aplica un test y luego se valoran los resultados del mismo.(10)

Relativos al hueso escafoides 8 artículos: Ghane MR et al., 2016 (11), Walenkamp MM et al., 2016 (12), Rhemrev SJ et al., 2010 (13), Unay K et al., 2009 (14), Roolker et al., 1998 (15), Esberger, 1994 (16), Freeland P et al., 1989 (17) y Waeckerle et al., 1987 (18).

**Tabla 3:** Escafoides.

<b>Autor y año</b>	<b>Sensibilidad %</b>	<b>Especificidad %</b>	<b>Valor predictivo positivo (VPP)</b>	<b>Valor predictivo negativo (VPN)</b>
<b>Ghane MR et al., 2016</b>	95.23%	74.07%	74.07%	95.23%
<b>Walenkamp MM et al., 2016</b>	90%	67%	-	-
<b>Rhemrev SJ et al., 2010</b> Agarre*	15%/92%*	98%/34%*	61%	85%
<b>Unay K et al., 2009</b> Pinza pulgar-índice/ dolor pronación	73%/79%	75%/58%	96%/82%	23%/54%
<b>Roolker et al., 1998</b>	24%	85%	57%	65%
<b>Esberger, 1994</b>	70.5%	21.8%	41.9%	48%
<b>Freeland P et al., 1989</b> Dolor ASB/dolor ST/ combinado*	90%/87%/100%*	40%/57%/40%*	18.75%*	100%*
<b>Waeckerle et al., 1987</b> Dolor ASB/ Dolor SCR/ dolor CLP	100%/100%/97.5%	98%/98%/97.5%	-	-

ASB: tabaquera anatómica; ST: tubérculo del escafoides; SCR: supinación contra resistencia; CLP: compresión longitudinal pulgar

Los estudios Ghane MR et al., 2016, Freeland P et al., 1989 y Waeckerle et al., 1987, incluyen en el examen físico la valoración del dolor de la tabaquera anatómica, dolor en el tubérculo del escafoides y el test de la

compresión longitudinal del pulgar. Estos estudios muestran la mayor sensibilidad y valor predictivo negativo.

Mientras que *Unay K et al., 2009*, tiene un valor predictivo positivo muy alto; es *Rhemrev SJ et al., 2010*, el que tiene mayor especificidad.

Por su parte *Roolker et al., 1998* y *Esberger, 1994* tienen resultados pobres en cuanto a todos los valores.

Relativos a la mandíbula 4 artículos: *Neiner J et al., 2016* (19), *Caputo et al., 2013* (5), *Schwab et al., 1998* (20) y *Alonso LL et al., 1995* (21).

**Tabla 4:** Mandíbula.

Autor y año	Sensibilidad %	Especificidad %	Valor predictivo positivo (VPP)	Valor predictivo negativo (VPN)
Neiner J et al., 2016	88.5%	95%	88.5%	95%
Caputo et al., 2013	95%	68%	77%	92%
Schwab et al., 1998	95%	67%	63%	96%
Alonso LL et al., 1995	95.7%	63.5%	66.2%	95.2%

Todos los estudios muestran valores muy altos en todos los parámetros, excepto en la especificidad y en el valor predictivo positivo los cuales solo mantienen resultados excepcionales en *Neiner J et al., 2016*.

Con respecto a la columna, incluyendo columna lumbar, torácica y cervical 7 artículos: *Inaba et al., 2015* (22), *Inaba et al., 2011* (23), *Langdon et al., 2010* (24), *Duane et al., 2007* (25), *Siminoski K et al., 2003* (26), *Hoffman JR et al., 2000* (27) y *Gonzalez RP et al., 1999* (28).

**Tabla 5:** Columna.

<b>Autor y año</b>	<b>Sensibilidad %</b>	<b>Especificidad %</b>	<b>Valor predictivo positivo (VPP)</b>	<b>Valor predictivo negativo (VPN)</b>
<b>Inaba et al., 2015</b> (TL)	78.4%	72.9%	21.4%	97.3%
<b>Inaba et al., 2011</b> (TL)	48.2%	84.9%	22.7%	94.7%
<b>Langdon et al., 2010</b> (TL) signo percusión/ supino	87.5%/81.25%	90%/93.3%	93%/95%	82%/75%
<b>Duane et al., 2007</b> (C)	76.9%	54.7%	15.5%	95.7%
<b>Siminoski et al., 2003</b> (L) ≤2 dedos	87%	47%	-	-
<b>Hoffman JR et al., 2000</b> (C)	99%	12.9%	2.7%	99.8%
<b>Gonzalez RP et al., 1999</b> (C)	91%	82.3%	7.3%	99.8%

TL: toraco-lumbar; C: cervical; L:lumbar

Encontramos resultados muy variados. Destacan los buenos resultados en todos los parámetros de *Langdon et al., 2010*. Los dos estudios más antiguos incluidos, ambos sobre las cervicales, presentan los mejores resultados en cuanto a sensibilidad y valor predictivo negativo. En cuanto a los estudios realizados por *Inaba et al.*, pese a tener un VPN muy alto, el resto de parámetros no son brillantes.

En cuanto a fracturas de pelvis 2 artículos: Ham SJ et al., 1996 (29) y Grant PT, 1990 (30).

**Tabla 6:** Pelvis.

<b>Autor y año</b>	<b>Sensibilidad %</b>	<b>Especificidad %</b>	<b>Valor predictivo positivo (VPP)</b>	<b>Valor predictivo negativo (VPN)</b>
<b>Ham SJ et al., 1996</b> Elevación pierna recta	90%	95%	95%	90%
<b>Grant PT, 1990</b> Compresión + distracción/ uno de los dos	27%/59%	79%/71%	67%/76%	40%/52%

Dentro de los 4 test realizados, cabe destacar los resultados del test de elevación de la pierna recta en *Ham SJ et al., 1996*, los cuales tienen valores muy altos. Los resultados del resto de test son pobres y no se reflejan en la tabla.

En Grant PT, 1990, los resultados son muy pobres tanto cuando en el test se valora el dolor a la compresión y distracción de forma combinada como cuando se valora de forma independiente.

A propósito de fracturas del codo tenemos 7 artículos: Jie KE et al., 2014 (31), Arundel et al., 2014 (32), Amiri H et al., 2012 (33), Appelboam A et al., 2008 (34), Darracq et al., 2008 (35), Lamprakis A et al., 2007 (36) y Hawksworth and Freeland ,1991 (37).

**Tabla 7:** Codo.

Autor y año	Sensibilidad %	Especificidad %	Valor predictivo positivo (VPP)	Valor predictivo negativo (VPN)
<b>Jie KE et al., 2014</b> *Extensión + punto doloroso	91%/98%*	43%/11%*	50%/44.8%	88%/87.5%
<b>Arundel et al., 2014</b>	84%	54%	-	-
<b>Amiri H et al., 2012</b>	90%	92%	56%	98%
<b>Appelboam A et al., 2008</b>	96.8%	48.5%	45.8%	97.2%
<b>Darracq et al., 2008</b> Overall ROM/punto doloroso	100%/100%	97%/67%	-	-
<b>Lamprakis A et al., 2007</b>	92%	61%	55%	93%
<b>Hawskworth et al., 1991</b>	90.7%	58.7%	72%	84.3%

Independientemente del test realizado todos tienen una alta sensibilidad con una media de 92%. La especificidad sin embargo baja bastante con una media de 64.8% y es más baja aún si solo tenemos en cuenta los estudios en los que el test únicamente consiste en la extensión de codo; media 53%. Sin embargo la capacidad de descartar una fractura cuando los pacientes dan un resultado negativo al examinar el rango de movimiento es de 94.5%. No tanto sin embargo el valor predictivo positivo que da valores más discretos.

La última selección se hace respecto a los artículos en los que se utiliza el diapasón o la percusión combinados con la auscultación para detectar la solución de continuidad del hueso ante una posible fractura. En este apartado tenemos 7 artículos: Jawad Z et al., 2012 (38), Fatima ST et al., 2012 (39), Moore MB, 2009 (40), Tiru et al., 2002 (41), Kazemi M et al., 2000 (42), Adams y Yarnold, 1997 (43) y Misurya et al., 1987 (44).

**Tabla 8:** Diapasón o percusión con auscultación.

Autor y año	Sensibilidad %	Especificidad %	Valor predictivo positivo (VPP)	Valor predictivo negativo (VPN)
Jawad Z et al., 2012	78%	82%	78%	82%
Fatima ST et al., 2012	79%	63%	88%	46%
Moore MB, 2009	83%	80%	66%	91%
Tiru et al., 2002	96%	86%	98%	75%
Kazemi M et al., 2000	86.8%	50%	89.2%	44.4%
Adams y Yarnold, 1997	79%	95.4%	93.8%	84%
Misurya et al., 1987	94%	-	-	-

Excepto en Fatima ST et al., 2012, los resultados del diapasón son buenos pero no excepcionales. La media de la sensibilidad de los estudios que utilizan el diapasón es del 84% y la de la especificidad del 69%. Por lo tanto es probable que los sujetos con fracturas tengan un test positivo con el diapasón.

Por lo general el VPP tiene mejores resultados que el VPN. Por lo tanto, con el uso de esta técnica de valoración, hay muchas probabilidades de tener una fractura cuando la valoración a través del diapasón o percusión es positiva.

## DISCUSIÓN

El objetivo de esta revisión fue revisar sistemáticamente los test clínicos para detectar fracturas y valorar la fiabilidad y la validez diagnóstica de estos test clínicos. Proporcionando así, datos concretos y objetivos que nos ayudasen a corroborar o descartar la utilidad de los test estudiados. Debido a la variabilidad tanto del tipo de fracturas como de los test aplicados, los estudios se discutieron en función de las agrupaciones realizadas anteriormente en el apartado de resultados.

### Escafoides

La mayoría de las fracturas del carpo suceden en el escafoides. El mecanismo lesional más frecuente consiste en una caída con dorsiflexión de muñeca y en varias ocasiones con extensión de codo.

Un inapropiado tratamiento en este tipo de fracturas puede llevarnos a necrosis avasculares, mala consolidación u osteoartritis. (11,45)

Estas fracturas son difíciles de diagnosticar, ya sea mediante el diagnóstico por imagen o tras un examen clínico. En ocasiones los pacientes con fractura de escafoides suelen tener radiografías normales, más aún si son valorados en fase aguda. Por ello muchas veces se catalogan como fracturas ocultas. Solo un 40% de los pacientes sometidos a radiografía tienen una fractura en la muñeca. Una regla validada de decisión clínica, que sea rápida de ejecutar, precisa y económica podría reducir substancialmente la cantidad de radiografías realizadas en estos casos. (11,12,45)

De los 8 artículos analizados en este apartado, 5 artículos usan la radiografía como prueba de referencia. Estos estudios coinciden en ser los más antiguos de los 8, mientras que los más recientes; *Ghane MR et al., 2016*, *Rhemrev SJ et al., 2010*, y *Unay K et al., 2009*, usan la resonancia magnética.

Actualmente no hay un acuerdo sobre cual deber ser la prueba de referencia para el diagnóstico de las fracturas de escafoides. Sin embargo,

la resonancia magnética es considerada la prueba más fiable y precisa de diagnosticarlas y por lo tanto algunos estudios sugieren que debe ser la primera opción cuando haya sospecha de fractura. Sin embargo, la resonancia magnética requiere más tiempo y dinero que otras técnicas que se han ido utilizando como la tomografía computarizada, la gammagrafía ósea o la radiografía. (11,45)

En cuanto a los test realizados para detectar las fracturas cabe destacar que son muy variados y diversos. No se puede extraer una regla general de todos ya que cada uno tiene un método diferente de valoración, pese a que sean similares entre algunos estudios.

Obviando a *Walenkamp MM et al., 2016* que estudia el método para detectar la fractura distal de radio; el resto de artículos valoran la forma para detectar la fractura de escafoides.

El resto de artículos, excepto *Roolker et al., 1998* que utiliza un aparato que produce vibraciones para detectar la fractura, llevan a cabo un examen físico para averiguar la presencia de la fractura. Encontramos exámenes físicos diferentes. Varios coinciden en el dolor de la tabaquera anatómica (ASB) a la palpación, dolor en el tubérculo del escafoides (ST) y compresión longitudinal/axial del pulgar (CLP). Otras pruebas que se han realizado han consistido en la flexo-extensión de muñeca y en la prono-supinación.

Los resultados son muy variados según la técnica utilizada. Los estudios que incluyen los test de dolor a la palpación, tienen buenos resultados; al igual que los que incluyen la compresión longitudinal del pulgar, excepto *Esberger, 1994*. Este último artículo tiene una sensibilidad del 70.5% y una especificidad del 21.8%. En el propio artículo se aclara que los resultados demuestran una fiabilidad muy baja y se desaconseja la utilización del test para detectar fracturas de escafoides.

En *Roolker et al., 1998*, se utiliza la vibración para tratar de detectar la fractura. Si el paciente refiere dolor a la vibración se considera un resultado positivo. Los resultados con esta técnica son muy pobres con una sensibilidad del 24%. Por lo tanto esta técnica se desaconseja para detectar

fracturas. La especificidad es algo más alta 85%, por lo que un resultado negativo de la prueba puede ayudarnos a pensar que el paciente no tiene una fractura.

En Rhemrev SJ et al., 2010, el test consiste en la flexo-extensión, prono-supinación y la fuerza de agarre. Se establece una regla combinando las tres pruebas que es llevada a cabo en todos los pacientes. Al combinarlas la sensibilidad baja mucho (15%) pero la especificidad (98%) y el valor predictivo negativo (85%) son muy altos. Por lo tanto es muy probable que tras esta regla si el resultado del test es negativo el paciente no tenga una fractura o que en los sujetos sanos el resultado de la prueba sea negativa. De las tres pruebas si analizamos la fuerza de agarre por si sola los resultados son muy positivos a la hora detectar fracturas en sujetos que la tienen; sensibilidad (92%).

Los estudios *Ghane MR et al., 2016*, *Freeland P et al., 1989* (17) y *Waeckerle et al., 1987*, se podrían agrupar por la similitud del examen físico realizado. Los tres se centran en los ya explicados conceptos de (ASB), (CT) y (CLP). Además son los estudios con un mejor resultado en cuanto a la sensibilidad, con una media de 98.3%. La especificidad no es tan alta y es en *Waeckerle et al., 1987* (17) en la que muestra mejores valores. En este último estudio también se menciona la importancia de la historia clínica y el mecanismo lesional a la hora de detectar la fractura; así como, la presencia de inflamación y dolor en la tabaquera anatómica o dolor a los movimientos de desviación radial y flexo-extensión de muñeca. En *Freeland P et al., 1989* (17), se destaca la importancia de combinar los test para obtener una mayor sensibilidad pese a perder especificidad. Finalmente en *Ghane MR et al., 2016*, se define que entre los test clínicos realizados el dolor en el tubérculo del escafoides (ST) es el más sensible y específico. Sin embargo recalca la importancia de combinar el examen físico con el diagnóstico por imagen para ser lo más preciso posible a la hora de detectar la fractura.



## **Mandíbula**

Las fracturas mandibulares son una de las fracturas más comunes vistas en los servicios de traumatología. Respecto a fracturas faciales, corresponden a un 40%-62% del total. Por lo tanto, un simple test de screening, como es el "Tongue Blade Test" o test pala-lengua (TBT), puede ayudar a los examinadores a establecer un diagnóstico rápidamente y reducir el tiempo de estancia de los pacientes. Además de reducir la cantidad de radiación ionizada a la que se ven sometidos los pacientes en las imágenes radiológicas. (5)

La forma de aplicar el test (TBT) difiere entre los estudios. En Alonso LL et al., 1995 y Caputo et al., 2013 la pala se gira tratando de romperla mientras el paciente aprieta y la sujeta con los dientes de uno u otro lado de la boca. En Neiner J et al., 2016 y Schwab et al., 1998, el examinador realiza una fuerza para tratar de sacar la pala de la boca mientras el paciente aprieta los dientes. La imposibilidad para romper la pala debido al dolor o no mantener la pala en la boca mientras se aplica la fuerza para sacarla es considerado un test positivo.

En cuanto a la prueba de referencia, los dos estudios más recientes utilizaron la tomografía computarizada; mientras que los dos más antiguos utilizaron la radiografía simple.

Independientemente del método utilizado para detectar las fracturas mandibulares, los resultados son muy favorables. La sensibilidad y el valor predictivo negativo son muy altos por lo que el test sirve tanto para detectar fracturas en pacientes que las tengan y para descartar fracturas en pacientes con un test negativo.

Pese a que los resultados son muy similares, en Neiner J et al., 2016 se recomienda utilizar su test ya que rompiendo la pala se pueden producir astillas que dañen tejidos bucales. Sin embargo, su test es menos objetivo ya que la fuerza que hay que aplicar para sacar la pala de la boca no se delimita o se describe como ligera-moderada, pudiendo variar mucho inter-examinador.

## Columna

En este apartado vamos a subdividir los estudios entre los que realizan un examen físico en la columna toraco-lumbar; *Inaba et al., 2015, Inaba et al., 2011, Langdon et al., 2010 y Siminoski K et al., 2003*. Con los que valoran la columna cervical; *Duane et al., 2007, Hoffman JR et al., 2000 y Gonzalez RP et al., 1999*.

En los traumas cerrados con afectación de la columna, es primordial un diagnóstico rápido y eficaz ya que las consecuencias pueden ser muy adversas. En las fracturas vertebrales; y sobre todo en inestables, un diagnóstico precoz es esencial para descartar lesiones medulares. Una lesión medular no hallada puede ser devastador para el futuro inmediato del paciente y susceptible de un empeoramiento iatrogénico. Sin embargo debido al coste, tiempo y los ya conocidos efectos adversos de la radiación, los pacientes deben ser cuidadosamente seleccionados a la hora de recibir un diagnóstico por imagen. (22)

Cada estudio tiene su método para valorar la presencia de una posible fractura en la columna. Aun así varios coinciden a la hora de valorar la presencia de un déficit neurológico, dolor, molestia o deformidad a la palpación de la región valorada. Además varios subrayan la importancia de valorar el nivel de alerta mediante la Escala de Coma de Glasgow (GCS), medir la intoxicación del paciente así como el mecanismo de producción de la posible lesión.

No hemos encontrado una prueba de referencia común en todos los estudios. Los que tratan sobre la columna toraco-lumbar (CTL) suelen utilizar la tomografía computarizada (TC) o la imagen por resonancia magnética (IRM). La evidencia defiende el uso de la TC como prueba de referencia en estos casos.(22)

En los casos de fracturas vertebrales por osteoporosis, es complicado diferenciar entre fracturas agudas y las ya cicatrizadas en radiografías simples. El diagnóstico de fracturas sintomáticas se debe realizar con la IRM. (24)

Los estudios que valoran la columna cervical (CC) utilizan más la radiografía pese haber un estudio con tomografía computarizada. Esto se puede deber a que los artículos son más antiguos. La TC es una técnica más apropiada para detectar este tipo de lesiones, las radiografías simples pueden ser imprecisas en muchos casos. (25)

En cuanto a la región cervical, *Hoffman JR et al., 2000* confirma la validez de un instrumento de diagnóstico basado en 5 criterios: ausencia de dolor cervical, ausencia de un déficit neurológico, nivel de alerta normal, no intoxicación y ausencia de lesión dolorosa o molesta. Los pacientes que cumplen los 5 criterios son considerados de bajo riesgo y no requieren diagnóstico por imagen. La aplicación de este protocolo de valoración alcanza una sensibilidad del 99% y del 100% en lesiones clínicamente significativas. Debido al buen resultado del valor predictivo negativo 99.8%, se puede afirmar que los pacientes con un trauma cervical cerrado en los que el protocolo de valoración de un resultado negativo, no tendrá una fractura cervical.

La habilidad de los examinadores para aplicar el protocolo con una sensibilidad tan alta y sin un entrenamiento previo exhaustivo, sugiere que dicho protocolo puede ser enseñado sin un gasto económico o humano importante. Además se decidió que cada criterio de valoración no sería definido explícitamente. Se cree que la presencia de intoxicación, molestias cervicales o el nivel de alerta son mejor valorados según el juicio clínico individual en lugar de seguir un criterio uniforme. (27)

Los dos estudios restantes sobre la columna cervical tienen resultados dispares.

En *Duane et al., 2007*, pese a seguir los pasos de *Hoffman Jr et al., 2000*, los resultados no son tan satisfactorios. Cabe destacar el VPN que sigue siendo muy positivo, 95.7%. El estudio recurre a dos razones con las que justificar sus resultados. La primera es debido a que utiliza la tomografía computarizada como prueba de referencia, técnica más precisa que la radiografía simple y por lo tanto con más posibilidades de encontrar fracturas potencialmente perdidas con la radiografía. La segunda, los examinadores no son instruidos para focalizarse exclusivamente en la

valoración cervical y por ende se lleva a cabo una valoración más general.(25)

Por el contrario, en Gonzalez *RP et al., 1999* los resultados son muy positivos y la metodología se asemeja mucho a la seguida y validada en *Hoffman JR et al., 2000*. Cabe resaltar que este estudio es llevado a cabo con muchos menos pacientes y que la prueba de referencia también es una radiografía. En este estudio las lesiones cervicales asintomáticas en un paciente despierto y que esté alerta, son consideradas estables y de bajo riesgo de lesión medular. Los resultados demuestran que los pacientes que han sufrido un trauma cervical cerrado con un  $GCS \geq 13$  y un examen cervical negativo tienen una probabilidad muy baja de tener una fractura cervical, además de tener una mejor sensibilidad que la radiografía lateral de la columna cervical. Finalmente destaca que tener niveles elevados de etanol no es una contraindicación a la hora de realizar un examen físico cervical. (28)

En cuanto a la columna toraco-lumbar, *Inaba et al., 2011 y 2015* realizaron dos estudios. El segundo fue más completo y tuvo mejores resultados. En él se consideraba un resultado positivo la presencia de una de las siguientes pruebas: dolor o molestia a la palpación, deformidad o un déficit neurológico. Además se determinaba la edad y el mecanismo de producción de la lesión. Los individuos  $\geq 60$  años y con un mecanismo de alto riesgo tenían una sensibilidad del 98.9%. Sin embargo, un 21.6% de los pacientes con lesiones clínicamente significativas tuvieron un examen físico negativo. Por ello el examen clínico por si solo es insuficiente para determinar la necesidad de un diagnóstico por imagen en pacientes con riesgo de fractura toraco-lumbar.

Finalmente mencionar los dos últimos estudios, *Langdon et al., 2010 y Siminoski K et al., 2003* los cuales usan métodos de valoración totalmente diferentes a los anteriores.

En *Langdon et al., 2010*, se utilizan dos pruebas clínicas para predecir fracturas por compresión de vértebras osteoporóticas sintomáticas. Las pruebas consisten en:

- Valorar el dolor a la percusión de las espinosas con la mano en puño mientras la otra está plana en la columna. "Closed-fist percussion test".
- Incapacidad de colocarse decúbito supino.

Ambas tienen muy buenos resultados en todos los parámetros, pero son menos fiables en la columna lumbar baja. Esto se puede deber al mayor grosor de los tejidos blandos en la primera prueba. En la segunda, la lordosis fisiológica mantendría la fractura asintomática.

Estos signos son indicadores fiables y se deberían utilizar para racionalizar los pacientes referidos a una resonancia magnética.(24)

En *Siminoski K et al., 2003* se determina la fiabilidad de utilizar la distancia costilla-pelvis (DCP) para detectar fracturas lumbares. La distancia es medida con la anchura de los dedos, y es el mismo examinador el que determina la distancia en todos los pacientes. La razón de verosimilitud aumenta de 0.1 cuando la DCP es de cuatro dedos a 11.5 cuando no entra ningún dedo. Cuando la razón de verosimilitud de un test positivo es mayor de 10 la utilidad clínica de la prueba es altamente relevante. Los porcentajes mostrados en la tabla hacen referencia a una distancia  $\leq 2$  dedos, con una sensibilidad alta del 87%. La DCP era mayor en pacientes con fracturas torácicas en comparación con los pacientes con fracturas lumbares. Los pacientes con menor distancia eran los que habían sufrido ambas fracturas.

Para futuros estudios se recomienda que la medición se realice con regla y no con los dedos, muestras que incluyan al género masculino y estudios con una menor prevalencia de fracturas. Además se debería medir la pérdida de altura y el grado de cifosis torácica en pacientes con riesgo de osteoporosis.(26)

## Pelvis

En este apartado se ha recabado menos información al consistir únicamente en dos artículos y poco recientes. Ambos realizan un examen físico para detectar una fractura de pelvis. El diagnóstico precoz de las fracturas de pelvis es vital para la supervivencia del paciente. (30)

En *Ham SJ et al., 1996*, se realizan 4 test pero solo el test que consiste en levantar la pierna activamente tiene resultados destacables. El resto de los test consisten en la aparición de dolor ante la presión en el pubis y compresión o distracción en los iliacos.

En *Grant PT, 1990*, el test es similar a dos del estudio anterior. Este test consiste en hacer una compresión o distracción en la pelvis variando las tomas. La presencia de dolor será indicativo de un test positivo.

Ambos estudios utilizan la radiografía como prueba de referencia. Tras una sospecha de lesión pélvica y en personas de avanzada edad principalmente, un examen radiológico está totalmente justificado. (29)

Cabe destacar los resultados del test de elevación de la pierna recta en *Ham SJ et al., 1996*, los cuales tienen valores muy altos. Se cree que es debido a que durante el test varios músculos que se originan o insertan en la pelvis son contraídos produciendo síntomas en el paciente. Por otro lado en un 68% de los pacientes sin fractura se vio un test negativo pero dolor al ejecutarlo. Por lo tanto la sola presencia de dolor no se puede considerar como un valor predictivo.

En *Grant PT, 1990*, los resultados son muy pobres. Por lo tanto, el test no tiene validez diagnóstica y como el propio autor recalca, debe ser descartado en la práctica clínica.

## Codo

Las lesiones de codo son relativamente frecuentes, corresponden a un 2%-3% de los casos en urgencias. La mayoría de los pacientes son sometidos a radiografías pese que solo un 30%-40% sufren una fractura. Para prevenir radiografías innecesarias en lesiones de extremidades, se suelen utilizar reglas de decisión clínica. Sin embargo, esta regla aún no existe para lesiones en el codo.(31)

Excepto en *Amiri H et al., 2012* y *Darracq et al., 2008* en los que se evalúa el rango de movimiento del codo en todas direcciones (flexión, extensión, supinación y pronación), en el resto de artículos el test consiste en realizar una extensión activa. La incapacidad de extender de forma activa el codo es considerado como un test positivo. En los dos estudios más recientes y en *Darracq et al., 2008*, se observa la combinación de la incapacidad de extender con puntos dolorosos en el codo.

Todos los estudios son comparados con la misma prueba de referencia; una radiografía.

Los estudios que se centran únicamente en la extensión tienen menor especificidad. Además casi todos los pacientes con fractura tienen cierta pérdida de movilidad en todas direcciones. Esta pérdida es más acusada en extensión. (33)

Según *Darracq et al., 2008* todos los pacientes identificados radiológicamente con una fractura tienen una disminución del rango en el codo y puntos dolorosos en el codo.

Según *Lamprakis A et al., 2007* el test de extensión de codo tiene una alta sensibilidad y defiende que los pacientes con test positivo deben ser referidos al hospital para llevar a cabo una radiografía. El test facilita el screening de los pacientes que deberían ser derivados al hospital, reduciendo en un 40% radiografías innecesarias y sus consiguientes costes.

*Jie KE et al*, 2014 defiende que hay que tener cuidado a la hora de descartar la radiografía en pacientes con extensión completa ya que en su estudio no se hubiesen detectado un 12% de las fracturas. En este estudio se explican varias limitaciones; entre otras, se perdió un 22% de la muestra en el seguimiento.

En Appelboom A et al., 2008 se muestra una muy baja LH- o razón de verosimilitud para un test negativo lo que confirma que este test es influyente a la hora de descartar fracturas en adultos.

### **Diapasón o auscultar más estetoscopio**

En este último apartado se valoran las fracturas a través de instrumentos de valoración. Cinco de los siete artículos, *Jawad Z et al.*, 2012, *Fatima ST et al.*, 2012, *Moore MB*, 2009(40), *Kazemi M et al.*, 2000 y *Misurya et al.*, 1987, usan la técnica del diapasón y la auscultación mediante un estetoscopio. Los dos artículos restantes; *Tiru et al.*, 2002 y *Adams y Yarnold*, 1997 valoran la fractura de cadera mediante la percusión sin diapasón pero con el estetoscopio para auscultar.

Uno de los test más comunes a la hora de detectar fracturas y que ya se propuso hace 60 años consiste en el uso del diapasón. Se han desarrollado dos métodos para detectar las fracturas a través del diapasón.

La primera consiste en hacer vibrar el diapasón y colocarlo encima o muy próximo al foco de fractura. Debido a que el periostio está altamente inervado, la vibración mecánica estimulará el periostio superficial produciendo dolor. Esta técnica está más en desuso que la segunda y es menos fiable.

En la segunda hacemos vibrar el diapasón y lo colocamos en una prominencia ósea distal a la fractura. La fractura se detectará por la interrupción de la conducción del sonido por el hueso, para ello utilizaremos un estetoscopio y se puede comparar con el lado contralateral. (46)



Se trata de un test fácil y rápido de aplicar, indoloro, y no supone un coste económico grande.(40)

El test del diapasón es efectivo únicamente cuando se ha perdido toda la integridad del hueso y cuando no hay un exceso de tejido rodeando la fractura.(3)

La prueba de referencia varía en función del hueso fracturado. En las fracturas de cadera, los últimos estudios destacan la imagen por resonancia magnética como la herramienta diagnóstica más fiable, la radiografía simple obtiene resultados inconclusos y la resolución del hueso trabecular osteoporótico es limitada. No hay evidencia sobre el uso de la tomografía computarizada en fracturas ocultas de cadera. (38)

Sin embargo estudios más antiguos, *Tiru et al., 2002* y *Adams y Yarnold, 1997*, utilizaron la radiografía simple para la detección de la fractura de cadera.

El resto de estudios utilizaron la radiografía simple ya que es el método más utilizado para la detección de fracturas en huesos largos.

En *Moore MB et al., 2009*, se valora el uso del diapasón en diferentes tipos de fracturas. En las fracturas completas el sonido es claramente interrumpido. En las transversales, el espacio creado por la fractura es suficiente para disminuir el sonido que produce el diapasón. En las avulsiones se produce una lesión ósea pero el sonido no suele verse afectado. En las fracturas impactadas, el sonido no se pierde, obteniendo resultados negativos. Además la inflamación del tejido o articulación puede alterar el resultado de la prueba por lo que el test se debe tratar de realizar antes de que esto ocurra. En caso de inflamación el diapasón se coloca encima del edema. Pese a tener resultados aceptables no debe utilizarse como única técnica de diagnóstico cuando hay sospecha de fractura. (40)

En *Kazemi M et al., 2000*, se comparan dos tipos de diapasón con diferente frecuencia. Primero se lleva a cabo con uno de 128Hz y luego con el de 256H. En el estudio se destaca que este proceso debería haber sido aleatorizado. Ambos diagnosticaron el mismo número de fracturas pero no en los mismos pacientes. En este estudio se compara con estudios previos

como los de *Bache and Cross 1984*, en el que se trata de diagnosticar una fractura del cuello femoral con una sensibilidad del 91% y una especificidad del 82%. También se compara con *Lesho et al., 1997*, el cual evalúa la parte anterior de la tibia con una sensibilidad del 75% y una especificidad del 67%. Finalmente cabe destacar que el bajo resultado de la especificidad puede deberse a los pocos sujetos con ausencia de fractura incluidos en el estudio.

Misurya et al., 1987 es el estudio más antiguo de los especificados y la estadística no está completamente definida como en los estudios más recientes. Se muestra una sensibilidad alta del 94%. En este estudio se especifica la técnica de diagnóstico, en la cual en función de la colocación del diapasón podemos diferenciar entre una fractura de cuello de fémur, diáfisis de fémur o rama pubiana. En comparación con otros estudios más antiguos; *Peltier 1977*, *Carter 1981* y *Colwell and Berg 1958*, utilizaron la percusión en lugar de utilizar el diapasón. Sin embargo, en este estudio se mostró que el test del diapasón es más sencillo de realizar y más fiable.

Jawad Z et al., 2012, resultó ser el estudio más reciente. La técnica de valoración de la fractura es más objetiva en este estudio. Esto se debe a la utilización de un estetoscopio electrónico el cual recibe y guarda la onda del sonido transmitido por el hueso. En el resto de estudios era el examinador el que determinaba la pérdida del sonido a través del hueso con un estetoscopio convencional. En este caso los datos se analizaban mediante un programa que convertía el sonido en una reproducción de audio y otra reproducción visual proyectada en un monitor. De esta manera se podía calcular el ratio de la amplitud del sonido, dividiendo la amplitud del lado sintomático (baja) con la del lado asintomático (alto). Pese a que los resultados no fueron extraordinarios y la muestra fue pequeña, cabe destacar la importancia del estudio por mostrar un método para objetivar y analizar la transmisión del sonido en fracturas, reduciendo la subjetividad de este tipo de test diagnóstico. (38)

Los resultados mostrados en *Fatima ST et al., 2012* son los más pobres de la tabla. El test del diapasón es menos fiable en fracturas por estrés

comparado con el resto de fracturas. Esto se puede deber a que no hay una pérdida de continuidad total del hueso y la superficie ósea está relativamente intacta. En estos casos el tiempo de espera hasta que se hace la prueba del diapasón también es muy importante ya que el hueso ha podido empezar a osificar y en un callo de fractura en el que ya ha empezado el proceso de mineralización el diapasón puede ser más inexacto. (46)

*Tiru et al., 2002* y *Adams y Yarnold, 1997*, son los dos únicos estudios incluidos en la revisión en los que se utiliza la percusión en lugar del diapasón. Ambos valoran la fractura de cadera; y ambos utilizan exactamente el mismo test. Se realiza la percusión en la rótula y se ausculta la sínfisis pubiana valorando la posible disminución del tono en la comparación con el lado sano. Mientras que *Tiru et al., 2002* obtuvo una sensibilidad mayor, fue *Adams y Yarnold, 1997* el que obtuvo mayor especificidad. Los dos obtuvieron un VPP muy alto por lo que la probabilidad de tener la fractura en caso de obtener un test positivo es muy alta. Además en el estudio más reciente también se calculó la razón de verosimilitud; la cual tenía un "buen" impacto 6.73 para un test positivo pero "mala" 0.75 para un test negativo.

La presencia de lesiones en tejidos blandos, edemas articulares o lesiones no traumáticas como tumores o quistes pueden interferir en la valoración dificultando la transmisión de sonido. (43)

En *Tiru et al., 2002*, aparece uno pequeño pero significativo porcentaje de falsos negativos (3.4%) y de falsos positivos (1.7%), por ello se delibera que el test de percusión y auscultación es en el mejor de los casos una técnica clínica que ayude al diagnóstico de las fracturas de cadera pero nunca como sustitutivo de una imagen radiológica. Esta idea se plasma en los dos estudios, no se trata de una técnica supletoria sino como un test más para llegar a un diagnóstico eficaz. (41,43)

## **Limitaciones del estudio**

Debido a la complejidad y extensión de la búsqueda realizada, solo se utilizó la base de datos PUBMED. Esta base de datos nos permitió agrupar y combinar los términos para realizar la búsqueda de la manera más precisa.

De los 50 artículos seleccionados en base a la lectura del título y el abstract, hubo 10 de los que no fue posible conseguir el artículo completo, ya que con las herramientas que disponíamos no se pudieron descargar.

Ciertos artículos relacionados con el tema de estudio no aparecieron en la búsqueda inicial y por lo tanto no se pudieron incluir. Sin embargo, algunos de ellos aparecen en la discusión de nuestros artículos y se compararon los resultados.

La revisión es llevada a cabo por una única persona, por lo tanto el riesgo de sesgo en la evaluación de la calidad metodológica es mayor.

## **Sugerencias para futuros estudios**

Se sugiere seguir investigando en la aplicación de todo tipo de test clínicos para detectar fracturas. De este modo se podrán realizar revisiones más amplias y habrá mas resultados para comparar la validez de estos test diagnósticos.

En cuanto a las revisiones sistemáticas en concreto; se pueden acotar los criterios de búsqueda a un tipo de test clínico específico que disponga de varios artículos para ser analizados. De esta manera se obtendrá una información más detallada y precisa del test clínico en cuestión.

## **Implicaciones clínicas**

En cuanto a la implicación clínica que puede suponer esta revisión, el lector puede sacar un juicio propio de que test clínico puede ser útil y válido a

través de los resultados mostrados en las tablas. En las conclusiones se mostrarán las decisiones sobre la aplicabilidad de cada test clínico.

En cuanto a las fracturas de escafoides, el examen físico debe incluir la valoración del dolor en la tabaquera anatómica, dolor en el tubérculo del escafoides y el test de la compresión longitudinal axial del pulgar. Para conseguir mayor precisión es necesario valorar la prono-supinación y la flexo-extensión de muñeca, además del mecanismo lesional. Estos test independientemente no sirven por lo que hay que realizar la valoración completa. Aun así, siempre que haya duda se debe corroborar con un diagnóstico por imagen. Cabe destacar que la valoración mediante el vibrador no es fiable.

Este estudio muestra que los dos test diagnósticos, descritos en la discusión (tirar y romper la pala), para detectar fracturas mandibulares son validos. Ciertamente es que uno es más objetivo que el otro y por lo tanto se puede utilizar como referencia. Esto ayudará a reducir la cantidad de TC realizadas, principalmente en casos de bajo riesgo.

Los resultados del presente estudio indican que el examen físico es insuficiente para obviar la necesidad de un diagnóstico por imagen en fracturas de columna toraco-lumbar. Incluir la edad y el mecanismo lesional en el examen mejora los resultados pero no es suficiente.

Los test aplicados en *Langdon et al., 2010* y *Siminoski K et al., 2003*, pueden complementar la valoración de las fracturas osteoporóticas por compresión y racionalizar los pacientes derivados a IRM, pero no se deben utilizar como única decisión diagnóstica.

Por su parte, en la columna cervical se define el método de decisión clínica basado en 5 criterios para identificar pacientes que han sufrido un trauma cervical con una probabilidad muy baja de tener una fractura cervical.

Los resultados obtenidos para detectar la fractura de pelvis parecen indicar que los test aplicados no tiene validez diagnóstica y por lo tanto no se deben utilizar en la valoración de este tipo de fracturas. Únicamente el test de elevar la pierna recta de forma activa tiene resultados reseñables.

Los resultados del estudio en relación al diagnóstico de las fracturas de codo generan una decisión unánime. El test de extensión de codo es válido para descartar o identificar posibles fracturas a este nivel. Sin embargo, conviene completar el examen físico con la valoración del rango de movimiento y la presencia de puntos de dolor no aumenta la precisión diagnóstica.

El test del diapasón es útil pero siempre como complemento y no como herramienta única para establecer una decisión diagnóstica. Pese a que en el estudio más reciente se consigue objetivar mucho la valoración de la fractura con un estetoscopio electrónico, es necesario un diagnóstico por imagen para confirmar la fractura. Lo mismo ocurre con los test de percusión, complementan la valoración de la fractura pero no se deben utilizar como sustitutivos del diagnóstico por imagen.

Por otro lado se descarta totalmente el uso del diapasón en fracturas por estrés.

## **CONCLUSIONES**

En la conclusión, al igual que se ha hecho en el resto del estudio, se mostraran las valoraciones finales por subgrupos. De este modo se especificará la utilidad clínica de los test estudiados en función de su validez diagnóstica.

Debido a la heterogeneidad de las valoraciones realizadas incluso dentro de un mismo subgrupo, sacar resultados concluyentes en complejo.

La conclusión principal del estudio es que los test para detectar fracturas de escafoides, mandíbula, codo o columna cervical son válidos. Sin embargo, los test para detectar fracturas de columna toraco-lumbar, excepto el test para vértebras osteoporóticas, y pelvis no son válidos. Los test con la herramienta diapasón, se deben seguir investigando, muestran cierta validez pero no se deben utilizar como método diagnóstico único.

La conclusión de los objetivos secundarios, desglosados por regiones del cuerpo es la siguiente:

- Pie-tobillo. No se encuentran artículos sobre test clínicos que valoren fracturas en esta región.
- Pierna-rodilla. Los test con diapasón, en huesos largos, deben seguir siendo investigados para obtener resultados más concluyentes. Se descarta totalmente en fracturas por estrés. Los test con percusión pese a tener buenos resultados, la pobre metodología y antigüedad del estudio nos hace sospechar de posibles sesgos por lo que también se descarta su aplicación.
- Muslo-pelvis. Los test clínicos para detectar fracturas de pelvis no tienen validez diagnóstica.
- Columna lumbo-torácica. Excepto el test para vértebras osteoporóticas, no hay examen físico o test clínico para detectar fracturas en esta región.

- Columna cervical. El método basado en los 5 criterios para detectar fracturas a nivel cervical tiene validez diagnóstica para ser aplicado.
- Región cráneo-facial. Los dos tipos de test clínicos descritos en la revisión son válidos para detectar fracturas mandibulares.
- Escápula-brazo. No se encuentran artículos sobre test clínicos que valoren fracturas en esta región.
- Codo-antebrazo. El test de extensión de codo es válido a la hora de diagnosticar fracturas a este nivel.
- Muñeca-mano. El examen físico descrito anteriormente tiene validez para diagnosticar las fracturas de escafoides.



## **PALABRAS CLAVE**

Fractures, bone, sensitivity and specificity, validity, physical examination, clinical test, diagnostic tests, diagnosis.

## **DICCIONARIO**

*ASB: tabaquera anatómica*

*ST: tubérculo del escafoides*

*SCR: supinación contra resistencia*

*CLP: compresión longitudinal pulgar*

TBT: Tongue Blade Test

*TL: toraco-lumbar*

*C: cervical*

*L: lumbar*

DCP: distancia costilla-pelvis

TC: Tomografía computarizada

IRM: Imágen por resonancia magnética

## **BIBLIOGRAFÍA**

1. Ruiz del pino MJ. Fracturas: conceptos generales y tratamiento. 2004;1-27.
2. Schünke, Schulte, Schumacher, Voll W. PROMETHEUS Texto y atlas de anatomía, Tomo 1: Anatomía general y aparato locomotor. Biblioteca Nacional. 2010. 100-128 p.
3. Toney CM, Games KE, Winkelmann ZK, Eberman LE. Using tuning-fork tests in diagnosing fractures. J Athl Train. 2016;51(6):498-9.
4. Gallo Vallejo FJ, Ruiz VG. Diagnóstico. Estudio radiológico. Ecografía, tomografía computarizada y resonancia magnética. Aten Primaria [Internet]. 2014;46(SUPPL 1):21-8. Available from: [http://dx.doi.org/10.1016/S0212-6567\(14\)70040-X](http://dx.doi.org/10.1016/S0212-6567(14)70040-X)
5. Caputo ND, Raja A, Shields C, Menke N. Re-evaluating the diagnostic accuracy of the tongue blade test: Still useful as a screening tool for mandibular fractures? J Emerg Med [Internet]. 2013;45(1):8-12. Available from: <http://dx.doi.org/10.1016/j.jemermed.2012.11.078>
6. Schneiders AG, Sullivan SJ, Hendrick PA, Hones BDGM, McMaster AR,

- Sugden BA, et al. The Ability of Clinical Tests to Diagnose Stress Fractures: A Systematic Review and Meta-analysis. J Orthop Sport Phys Ther [Internet]. 2012;42(9):760–71. Available from: <http://www.jospt.org/doi/10.2519/jospt.2012.4000>
7. Silva Fuente-Alba C, Molina Villagra M. Likelihood ratio (razón de verosimilitud): definición y aplicación en Radiología. Rev Argentina Radiol [Internet]. 2017;81(3):204–8. Available from: <http://dx.doi.org/10.1016/j.rard.2016.11.002>
  8. Moher D, Liberati A, Tetzlaff J, Altman DG, Altman D, Antes G, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. 2009;6(7).
  9. Ciapponi A. QUADAS-2 : instrumento para la evaluación de la calidad de estudios de precisión diagnóstica QUADAS-2 : an instrument for the evaluation of the quality of diagnostic precision studies. Evidencia [Internet]. 2015;22–6. Available from: <http://www.foroaps.org/files/64fe85009abba8c506e903adf90dbc17.pdf>
  10. Manterola C, Otzen T. Estudios Observacionales: Los Diseños Utilizados con Mayor Frecuencia en Investigación Clínica. Int J Morphol [Internet]. 2014;32(2):634–45. Available from: [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0717-95022014000200042&lng=en&nrm=iso&tlng=en](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-95022014000200042&lng=en&nrm=iso&tlng=en)
  11. Ghane MR, Rezaee-zavareh MS, Kazem M, Dehghani V. How Trustworthy Are Clinical Examinations and Plain Radiographs for Diagnosis of Scaphoid Fractures ? 2016;21(5).
  12. Rosenwasser MMJWMP, Schep JCGNWL. A multicentre cross - sectional study to examine physicians ' ability to rule out a distal radius fracture based on clinical findings. 2016;185–90.
  13. Rhemrev SJ, Beeres FJP, Leerdam RH Van, Hogervorst M, Ring D. Clinical prediction rule for suspected scaphoid fractures A Prospective Cohort Study. Injury [Internet]. 2010;41(10):1026–30. Available from: <http://dx.doi.org/10.1016/j.injury.2010.03.029>
  14. Unay K, Gokcen B, Ozkan K, Poyanli O, Eceviz E. Examination tests predictive of bone injury in patients with clinically suspected occult scaphoid fracture. 2009;40(February 2008):1265–8.

15. Roolker L, Tiel-Van Buul MMC, Broekhuizen TH. Is intrasound vibration useful in the diagnosis of occult scaphoid fractures? *J Hand Surg Am*. 1998;23(2):229–32.
16. Volume E. WHAT VALUE THE SCAPHOID COMPRESSION TEST ? 1994;
17. Freeland P. Scaphoid tubercle tenderness: a better indicator of scaphoid fractures? *Arch Emerg Med* [Internet]. 1989;6(1):46–50. Available from:  
<http://www.ncbi.nlm.nih.gov/pubmed/2712988>  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1285557>
18. Waeckerle JF, City K. A Prospective Study Identifying the Sensitivity of Radiographic Findings and the Efficacy of Clinical Findings in Carpal Navicular Fractures. 1987;(July):733–7.
19. Neiner J, Free R, Nathan C, Caldito G, Moore-medlin T. Tongue Blade Bite Test Predicts Mandible Fractures. 2016;1–4.
20. Schwab RA, Genners K, Robinson WA. Clinical predictors of mandibular fractures. *Am J Emerg Med* [Internet]. 1998;16(3):304–5. Available from:  
<http://www.sciencedirect.com/science/article/pii/S0735675798901085>
21. Academic MIN, Settings C. Original Contributions. 2008;xx(x):297–304.
22. Inaba K, Nosanov L, Menaker J, Bosarge P, Williams L, Turay D, et al. Prospective derivation of a clinical decision rule for thoracolumbar spine evaluation after blunt trauma : An American Association for the Surgery of Trauma Multi-Institutional Trials Group Study. 2015;78(3):459–67.
23. Inaba K, Dubose JJ, Barmparas G, Barbarino R, Reddy S, Talving P, et al. Clinical Examination Is Insufficient to Rule Out Thoracolumbar. 2011;70(1).
24. Langdon J, Way A, Heaton S, Bernard J, Molloy S. Vertebral compression fractures - New clinical signs to aid diagnosis. *Ann R Coll Surg Engl*. 2010;92(2):163–6.
25. Duane TM, Dechert T, Wolfe LG, Aboutanos MB, Malhotra AK, Ivatury RR. Clinical examination and its reliability in identifying cervical spine fractures. *J Trauma - Inj Infect Crit Care*. 2007;62(6):1405–8.

26. Siminoski K, Warshawski RS, Jen H, Lee KC. Accuracy of physical examination using the rib-pelvis distance for detection of lumbar vertebral fractures. *Am J Med.* 2003;115(3):233–6.
27. Centers P. The New England Journal of Medicine VALIDITY OF A SET OF CLINICAL CRITERIA TO RULE OUT INJURY TO THE CERVICAL SPINE IN PATIENTS WITH BLUNT TRAUMA. 2000;
28. Gonzalez RP, Fried PO, Bukhalo M, Holevar MR, Falimirski ME. Role of clinical examination in screening for blunt cervical spine injury11No competing interests declared. %. *J Am Coll Surg* [Internet]. 1999;189(2):152–7. Available from: <http://www.journalacs.org/article/S1072751599000654/fulltext>
29. Ham SJ, Walsumz ADI Van, Vierho PAM. Predictive value of the hip flexion test for fractures of the pelvis. 1996;27(8):543–4.
30. Grant PT. The diagnosis of pelvic fractures by ‘springing’. 1990;178–82.
31. Jie KE, Dam LF Van, Verhagen TF, Hammacher ER. Extension Test and Ossal Point Tenderness Cannot Accurately Exclude Significant Injury in Acute Elbow Trauma Selection of Participants. *Ann Emerg Med* [Internet]. 2014;64(1):74–8. Available from: <http://dx.doi.org/10.1016/j.annemergmed.2014.01.022>
32. Arundel D, Williams P, Townend W. Deriving the East Riding Elbow Rule (ER2): A maximally sensitive decision tool for elbow injury. *Emerg Med J.* 2014;31(5):380–3.
33. Amiri H, Vahdati SS, Fekri S, Zadegan SA, Shokoohi H, Rahimi-movaghar V. Does preservation of active range of motion after acute elbow injury rule out the need for radiography? *Akut dirsek travmasından sonra aktif eklem hareket açıklığının korunması.* 2012;18(6):479–82.
34. Appelboom A, Reuben AD, Benger JR, Beech F, Dutson J, Haig S, et al. Elbow extension test to rule out elbow fracture: multicentre, prospective validation and observational study of diagnostic accuracy in adults and children. :5–9.
35. Darracq MA, Vinson DR, Panacek EA. Preservation of active range of motion after acute elbow trauma predicts absence of elbow fracture. 2008;779–82.

36. Lamprakis A, Vlasits K, Siampou E, Lionis C. Can elbow-extension test be used as an alternative to radiographs in primary care ? 2009;4788.
37. Hawksworth CR, Freeland P. Inability to fully extend the injured elbow: an indicator of significant injury. Arch Emerg Med [Internet]. 1991;8(4):253–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1772537>  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1285793>
38. Jawad Z, Odumala A, Jones M. Objective sound wave amplitude measurement generated by a tuning fork . An analysis of its use as a diagnostic tool in suspected femoral neck fractures. Injury [Internet]. 2012;43(6):835–7. Available from: <http://dx.doi.org/10.1016/j.injury.2011.09.030>
39. Fatima ST, Jeilani A, Abbasi NZ, Aziz A, Khan K, Sheikh AS, et al. ORIGINAL ARTICLE VALIDATION OF TUNING FORK TEST IN STRESS FRACTURES AND ITS COMPARISON WITH RADIONUCLIDE BONE SCAN. 2012;24:180–2.
40. Moore MB. The use of a tuning fork and stethoscope to identify fractures. J Athl Train. 2009;44(3):272–4.
41. Mohan T, Goh S, Low B. Research forum abstracts: 150 Use of Percussion as a Screening Tool in the Diagnosis of Occult Hip Fractures. Ann Emerg Med. 1999;34(October):1999.
42. Kazemi, M., & Roscoe MW. Is the tuning fork test a reliable tool in detecting acute simple fractures? Int Sport J. 2000;4(2):1–8.
43. Adams SL, Yarnold PR. Clinical use of the patellar-pubic percussion sign in hip trauma. Am J Emerg Med. 1997;15(2):173–5.
44. Misurya RK, Khare A, Mallick A, Sural A, Vishwakarma GK. Use of tuning fork in diagnostic auscultation of fractures. Injury. 1987;18(1):63–4.
45. Burrows B, Moreira P, Murphy C, Sadi J, Walton DM. Scaphoid fractures: A higher order analysis of clinical tests and application of clinical reasoning strategies. Man Ther [Internet]. 2014;19(5):372–8. Available from: <http://dx.doi.org/10.1016/j.math.2014.05.007>
46. Mugunthan K, Doust J, Kurz B, Glasziou P. Is there sufficient evidence for tuning fork tests in diagnosing fractures? A systematic review. BMJ Open. 2014;4(8):1–5.

## **ANEXO I**

(((((((((("Fractures, Bone"[MeSH Terms]) OR "Mandibular Fractures"[MeSH Terms] OR "fractures") **AND**  
("Sensitivity and Specificity"[MeSH Terms] OR "sensitivity" OR "specificity" OR "validity" OR "reliability") **AND**  
("Physical examination"[Mesh] OR "Diagnostic Tests, Routine"[Mesh] OR "diagnostic tool" OR "screening tool" OR "test" OR "clinical test") **AND**  
("diagnostic accuracy" OR "diagnosis").

## ANEXO II

PROTOCOLO DE REVISIÓN									
<b>Fecha de la búsqueda: 27/02/2018</b> <b>Objetivo:</b> <b>Criterios de inclusión y exclusión de los estudios:</b>									
Base de datos	Límites	Estrat. de búsqueda (palabras)	Resultados Búsqueda	Excluidos Título / Abstract	Excluidos No texto completo	Excluidos duplicados	Excluidos Lectura profunda (Poner qué criterios no cumplen)	Incluidos en la revisión (Cumplen criterios de inclusión y exclusión (Incluir nombre artículo))	
PUBMED	Sin límites	TOTAL ((((((((("Fractures, Bone"[MeSH Terms]) OR "Mandibular Fractures"[MeSH Terms] OR "fractures") AND ("Sensitivity and Specificity"[MeSH Terms] OR "sensitivity" OR "specificity" OR "validity" OR "reliability") AND ("Physical examination"[MeSH] OR "Diagnostic Tests, Routine"[MeSH] OR "diagnostic tool" OR "screening tool" OR "test" OR "clinical test") AND ("diagnostic accuracy" OR "diagnosis"))	822	780 → 42	10 → 32	0 → 32	5 → 27 -Tsitsilonis et al., 2016 (No se especifica cuál es el test clínico o el examen físico que se realiza). -Bujize GA et al., 2012 (No especifica el gold-standard, en una comparación entre técnicas de imagen y examen físico para ver que prueba de referencia será la más adecuada). -Shlamovitz et al., 2009 (Por ser una revisión). -Gonzalez RP et al., 2002 (Los resultados de sensibilidad y especificidad mal definidos y calculados. No se determina el gold-standard con el que comparan el examen físico). -Knight P et al., 1998 (El test clínico aplicado requiere de aparataje específico poco común)	-Adams y Varnold, 1997 -Alonso LL et al., 1995 -Amiri H et al., 2012 -Appelboom A et al., 2008 -Caputo et al., 2013 -Barracq et al., 2008 -Duane et al., 2007 -Esberger, 1994 -Fatima ST et al., 2012 -Ghane MR et al., 2016 -Gonzalez RP et al., 1999 -Inaba et al., 2011 -Inaba et al., 2015 -Jawad Z et al., 2012 -Jie KE et al., 2014	N -Knight et al., 1998 -Lamprakis A et al., 2010 -Langdon et al., 2010 -P.T. Grant, 1990 -Rhemrev SJ et al., 2011 -Roolker et al., 1998 -S.J. Ham et al., 1996 -Schwab et al., 1998 -Siminoski K et al., 2002 -Tru et al., 2002 -Unay K et al., 2009 -Waedekerle et al., 1987 -Walenkamp MM et al.
Incluidos tras revisión de las referencias bibliográficas de artículos:									
-Scaphoid tubercle tenderness: a better indicator of scaphoid fractures? → Ghane MR et al., 2016 - Inability to fully extend the injured elbow: an indicator of significant injury. → Amiri H et al., 2012 -Validity of a set of clinical criteria to rule out injury to the cervical spine in patients with blunt trauma. → Duane et al., 2007 -Use of tuning fork in diagnostic auscultation of fractures. → Jawad et al., 2012 -Deriving the East Riding Elbow Rule (ER2): a maximally sensitive decision tool for elbow injury. → Jie KE et al., 2014									
Incluidos de trabajos no publicados, tesis o congresos:									
Incluidos tras contacto con expertos:									
-Tongue Blade Bite Test Predicts Mandible Fractures (Tutor: Miguel Malo) -Is the tuning fork test a reliable tool in detecting acute simple fractures? (Tutor: Miguel Malo) -The Use of a Tuning Fork and Stethoscope to Identify Fractures. (Tutor: Miguel Malo)									
								-Freeland P, 1989 -Hawthorn and Freeland, 1991 -Hoffman JR et al., 2000 -Misurva et al., 1987 -Arundel et al., 2014	+5 → 32
								-Neiner J et al., 2016 -Kazemi M et al., 2000 -Moore, 2009	+3 → 35